A MATHEMATICAL THEORY OF ADAPTIVE CONTROL PROCESSES

By Richard Bellman and Robert Kalaba

THE RAND CORPORATION, SANTA MONICA, CALIFORNIA

Communicated by Philip M. Morse, June 26, 1959

Introduction.—In the last few years, the mathematical theory of control processes has attracted a great deal of attention.[1–7] Problems of stability and optimization have been posed and resolved for deterministic and stochastic processes. Along with this intensive analytic effort has come a more searching examination of the relevance of existing mathematical models of the underlying physical processes.[8, 9] It is now recognized that questions of engineering feedback and industrial process control, design of automata, control of man-machine combinations, data-processing on a large scale, and many other activities as well, contain elements of uncertainty of various unconventional types which escape classical formulation and classical treatment. The processes that occur are "learning"[10–12] or "adaptive" processes in which it is required to act and learn simultaneously. Although some preliminary work has been done from the statistical side[13–15] the foundations for a general theory have not as yet been laid.

The purpose of this note is to indicate how a general mathematical framework can be constructed using the techniques of the theory of dynamic programming.[16] More detailed accounts of particular applications will appear in subsequent publications;[17, 18] see also Bellman and Kalaba,[19] and Bellman.[20]

Formulation.—In a deterministic process, the state of the system is specified by a point in phase space, p, and the transformations, $T(p, q)$, resulting from decisions q, yield points of the same nature. In stochastic processes, the state of the system is specified by coordinates of this type, and in addition, by probability distributions, "points" in a "phase space" of more general type. At a sufficient level of abstruction, these are projections of a general multi-stage decision process.

Introducing a further level of abstraction, we shall show how the same conceptual and analytic techniques, and computational methods as well, can be applied to the study of adaptive control processes.

Let the state of the system S be specified, as usual, by a point p in phase space, and, what is new, by an *information pattern*, P. This information pattern represents the information about the process that we wish to retain in order to guide our further actions. As a consequence of a decision, q, p is transformed into a new point p_1, given *a priori* by the transformation $p_1 = T_1(p, P; q, r)$, and P is transformed into a new information pattern, given *a priori* by the transformation $P_1 = T_2(p, P; q, r)$. Here r is a random vector variable, specified by an *a priori* probability distribution $dG(p, P; q, r)$, itself a part of the information pattern P.

After the decision has been made, p_1 may or may not be completely determined. Let us assume here, for the sake of simplicity, that these are known after the decision has been made. We shall suppose also that the transformations T_1 and T_2 are known, although in many adaptive processes, the determination of these functions, and even of the state and information patterns themselves, is an essential part of the problem.

With this format, let it be required to determine a sequence of decisions, $[q_1, q_2,$

$\ldots, q_N]$, which will minimize the expected value of a preassigned function of the terminal state, (p_N, P_N). The expected value is taken with respect to the set of *a priori* probability distributions.

Functional Equations.—Following the usual approach of dynamic programming, we introduce the sequence of functions

$$f_N(p, P) = \text{Min Exp } \phi(p_N, P_N) \tag{1}$$

Then the principle of optimality yields the functional equation

$$f_N(p, P) = \underset{q_1}{\text{Min}} \left[\int f_{N-1}(T_1(p, P; q, r), T_2(p, P; q, r))dG(p, P; q, r)\right] \tag{2}$$

for $N = 2, 3, \ldots$, with

$$f_1(p, P) = \underset{q}{\text{Min}} \left[\int \phi(T_1(p, P; q, r), T_2(p, P; q, r))dG(p, P; q, r)\right. \tag{3}$$

These relations can be used to establish the existence of optimal policies, and to obtain various structural characteristics, as in Bellman.[16]

Sufficient Statistics.—For analytic and computational purposes, it is desirable to replace the sequence $f_N(p, P)$, which may be a sequence of functionals due to the dependence upon P, by a sequence of functions. A number of devices, familiar in modern mathematical statistics, and collected under the heading of "sufficient statistics," permit this very important reduction. See[3, 21, 22] for applications of this method, and for further discussion of particular problems.

Discussion.—In view of the very uncertainty that exists as to how to treat uncertainty, there can be no pretense of ever erecting a definitive theory of adaptive processes. The erection of such a theory is itself a sequential learning process. What we wish to construct is a theory which is plausible, feasible and flexible, and which can be shown, under certain favorable circumstances, to converge over time, with probability one in any particular process, to the theory which we possess for treating deterministic and stochastic control processes. Matters of this nature will be discussed in detail in subsequent publications.

[1] Truxal, J., *Automatic Feedback Control System Synthesis* (McGraw-Hill, 1955).

[2] Laning, J., and R. Battin, *Random Processes in Automatic Control* (McGraw-Hill, 1956).

[3] Wiener, N., *Nonlinear Problems in Random Theory* (John Wiley and Sons, 1958).

[4] Bellman, R., "Dynamic Programming and the Computational Solution of Feedback Design Control Processes," AIEE Special Publication T-101, *Proc. Computers in Control Systems Conference*, 1957.

[5] Bellman, R., "Dynamic Programming and Stochastic Control Processes," *Information and Control*, **1**, 228–239 (1958).

[6] Bellman, R., and R. Kalaba, "On adaptive control processes," 1959 IRE National Convention Record (to appear).

[7] LaSalle, J. P., "Time Optimal Control Systems," these PROCEEDINGS, **45**, 573–577 (1959).

[8] Zadeh, L., "What Is Optimal?", *IRE Trans. on Information Theory*, **IT–4**, 3 (1958).

[9] McMillan, B., "Where Do We Stand?", *IRE Trans. on Information Theory*, **IT–3**, 173–4 (1957).

[10] Bush, R. R., and F. Mosteller, *Stochastic Models for Learning* (John Wiley and Sons, 1955).

[11] Bellman, R., T. E. Harris, and H. N. Shapiro, *Studies in Functional Equations Occurring in Decision Processes*, RM-878 (Rand, 1952).

[12] Flood, M. M., *On Game-Learning Theory and Some Decision-Making Experiments, Decision Processes* (John Wiley and Sons, 1954).

[13] Wald, H., *Sequential Analysis*, (John Wiley and Sons, 1947).

[14] Robbins, H., "Some Aspects of the Sequential Design of Experiments," *Bull. Amer. Math. Soc.*, **58**, 527–535 (1952).

[15] Thompson, W. R., "On the Theory of Apportionment," *Amer. Jour. Math.*, 450–457 (1935).

[16] Bellman, R., *Dynamic Programming* (Princeton University Press, 1957).

[17] Bellman, R., and R. Kalaba, "On Adaptive Control Processes," 1959 IRE National Convention Record, (to appear).

[18] Bellman, R., and R. Kalaba, *Dynamic Programming and Adaptive Processes I: Mathematical Foundation*, P-1416, (Rand, 1959).

[19] Bellman, R., and R. Kalaba, "On Communication Processes Involving Learning and Random Duration," 1958 IRE National Convention Record, Part 4, 1958.

[20] Bellman, R., "A Problem in the Sequential Design of Experiments," *Sankhya*, **16**, 221–229 (1956).

[21] Aoki, M., Ph.D. thesis, UCLA (to appear).

[22] Freimer, M., "A Dynamic Programming Approach to Adaptive Control Processes," Lincoln Laboratory (1959).

PARTITIONS IN HOMOGENEOUS, FINITE ABELIAN GROUPS

BY ECKFORD COHEN

UNIVERSITY OF TENNESSEE

Communicated by Jesse Douglas, June 15, 1959

Let r represent a positive integer. A (complex-valued) arithmetical function $f_r(n)$ is called *even* (mod r) if $f_r(n) = f_r((n, r))$ for all integers n, where (n, r) denotes the greatest common divisor of n and r. Since the integers (mod r) form an additive cyclic group, one may rephrase this definition as follows. Let C_r be a cyclic group of order r with elements α. A function $f_r(\alpha)$ defined in C_r will be termed even if $f_r(\alpha)$ is invariant under all automorphisms of C_r.

The notion of even function (mod r) has been extended to functions of several variables in a paper as yet unpublished. In that paper a function $f_r(n_1, \ldots, n_k)$ of k integral variables is defined to be (relatively) *even* (mod r) if $f_r(n_1, \ldots, n_k) = f_r((n_1, r), \ldots, (n_k, r))$ for all n_i; if, moreover, there exists an even function $F_r(n)$ such that $f_r(n_1, \ldots, n_k) = F_r((n_1, \ldots, n_k))$ for all n_i, then $f_r(n_1, \ldots, n_k)$ is defined to be *totally even* (mod r).

In the present note we restate in group-theoretical terminology some of the concepts and results of the above cited paper. In particular, one may reformulate the preceding definitions as follows. Let $G_r^{(k)}$ denote an (additive) abelian group, decomposable into a direct sum of k cyclic groups, $C_r^{(i)}$, each of order r,

$$G_r^{(k)} = C_r^{(1)} \oplus \ldots \oplus C_r^{(k)}. \qquad (1)$$

Let $f_r(\alpha)$ be a complex-valued function defined in $G_r^{(k)}$. We shall say that $f_r(\alpha)$ is *relatively even*, or more precisely, even relative to the decomposition (1), if $f_r(\alpha)$ is invariant under all automorphisms of $G_r^{(k)}$ which induce automorphisms in $C_r^{(i)}$, $i = 1, \ldots, k$. The function $f_r(\alpha)$ will be termed *totally even* if it is invariant under the totality of automorphisms of $G_r^{(k)}$. The equivalence of the latter definition with the definition of totally even function (mod r) stated above is a consequence of the fact that, in the homogeneous group $G_r^{(k)}$, each cyclic subgroup of order r is a direct summand.[1]